(Gil Kalai) The entropy/influence conjecture

16 August, 2007 in guest blog, math.CA, math.PR, question | Tags: boolean functions, entropy, Fourier transform, Gil Kalai, influence | by Gil Kalai

[*This post is authored by Gil Kalai, who has kindly "guest blogged" this week's "open problem of the week". –* T.]

The *entropy-influence conjecture* seeks to relate two somewhat different measures as to how a boolean function has concentrated Fourier coefficients, namely the total influence and the entropy.

We begin by defining the total influence. Let $\{-1,+1\}^n$ be the discrete cube, i.e. the set of ± 1 vectors $(x_1,\ldots,x_n)$ of length n. A *boolean function* is any function $f:\{-1,+1\}^n \to \{-1,+1\}$ from the discrete cube to {-1,+1}. One can think of such functions as "voting methods", which take the preferences of n voters (+1 for yes, -1 for no) as input and return a yes/no verdict as output. For instance, if n is odd, the "majority vote" function $\mathrm{sgn}(x_1+\ldots+x_n)$ returns +1 if there are more +1 variables than -1, or -1 otherwise, whereas if $1 \leq k \leq n$, the "k^{th} dictator" function returns the value x_k of the k^{th} variable.

We give the cube $\{-1,+1\}^n$ the uniform probability measure μ (thus we assume that the n voters vote randomly and independently). Given any boolean function f and any variable $1 \leq k \leq n$, define the *influence* $I_k(f)$ of the k^{th} variable to be the quantity

$$I_k(f) := \mu\{x \in \{-1,+1\}^n : f(\sigma_k(x)) \neq f(x)\}$$

where $\sigma_k(x)$ is the element of the cube formed by flipping the sign of the k^{th} variable. Informally, $I_k(f)$ measures the probability that the k^{th} voter could actually determine the outcome of an election; it is sometimes referred to as the Banzhaf power index. The *total influence* I(f) of f (also known as the *average sensitivity* and the *edge-boundary density*) is then defined as

$$I(f) := \sum_{k=1}^n I_k(f).$$

Thus for instance a dictator function has total influence 1, whereas majority vote has total influence comparable to $\sqrt{n}$. The influence can range between 0 (for constant functions +1, -1) and n (for the parity function $x_1 \ldots x_k$ or its negation). If f has mean zero (i.e. it is equal to +1 half of the time), then the *edge-isoperimetric inequality* asserts that $I(f) \geq 1$ (with equality if and only if there is a dictatorship), whilst the Kahn-Kalai-Linial (KKL) theorem asserts that $I_k(f) \gg \frac{\log n}{n}$ for some k. There is a result of Friedgut that if $I(f)$ is bounded by A (say) and $\varepsilon > 0$, then f is within a distance ε (in L^1 norm) of another boolean function g which only depends on $O_{A,\varepsilon}(1)$ of the variables (such functions are known as juntas).

Now we define the entropy. For this we recall the *Fourier-Walsh expansion*

$$f = \sum_{S \subset \{-1,+1\}^n} \hat{f}(S) U_S$$

of any real-valued function $f:\{-1,+1\}^n \to \mathbb{R}$, where $U_S:\{-1,+1\}^n \to \{-1,+1\}$ is the monomial boolean function

$$U_S(x_1,\ldots,x_n) = \prod\{x_i : i \in S\}$$

and $\hat{f}(S)$ is the Fourier-Walsh coefficient

$$\hat{f}(S) := \int_{\{-1,+1\}^n} f(x) U_S(x)\ d\mu(x).$$

Algebraically, the Fourier-Walsh expansion regards f as a polynomial combination of multilinear (square-free) monomials. It is of course a special case of the more general concept of a Fourier transform on an abelian group.

From Plancherel's theorem we have

$$\sum_{S \subset \{-1,+1\}^n} \hat{f}(S)^2 = 1.$$

The influence can be written similarly:

$$I(f) = \sum_{S \subset \{-1,+1\}^n} \hat{f}(S)^2 |S|.$$

Note that this (together with Plancherel's theorem) already implies the edge-isoperimetric inequality. The influence thus measures the average "height" of the spectrum, with each set S being viewed as having height |S|.

The *spectral entropy* E(f) of a boolean function f is defined as

$$E(f) = \sum_{S \subset \{-1,+1\}^n} \hat{f}(S)^2 \log_2 \frac{1}{\hat{f}(S)^2}$$

Roughly speaking, it measures the logarithm of the "width" of the Fourier transform of f; it is larger when the spectrum of f is "smeared out". From Jensen's inequality, the entropy is at least zero (with equality attained if and only if f is a monomial U_S for some S), and can be as large as n or so (this is the case for random functions, for instance).

Entropy/influence conjecture (Friedgut-Kalai, 1996) : We have $E(f) \ll I(f)$ (i.e. $E(f) \leq CI(f)$ for some absolute constant C.)

Comparing the Fourier-analytic formulae for the entropy and influence, we thus see that this conjecture asserts in some sense that if the Fourier coefficients are "smeared" then the Fourier coefficients are (on average) "high". One strange feature of this conjecture is worth noting: the spectral entropy is invariant under arbitrary $\mathbb{Z}/2\mathbb{Z}$ linear transformations, whereas the influence depends on a specific basis. Thus, the conjecture also implies that if Fourier coefficients are smeared, then they are high with respect to arbitrary bases.

On the other hand, both the entropy and influence are additive with respect to tensor products: if $f : \{-1,1\}^n \to \{-1,1\}$ and $g : \{-1,1\}^m \to \{-1,1\}$ are boolean functions, then the tensor product $f \otimes g : \{-1,1\}^{n+m} \to \{-1,1\}$ is such that $E(f \otimes g) = E(f) + E(g)$ and $I(f \otimes g) = I(f) + I(g)$. This fact is of course consistent with the conjecture.

The entropy/influence conjecture, if true, has several consequences:

1. **Influences under symmetry**. The motivation for the entropy/influence conjecture was to understand influences (and the related threshold behavior) under symmetry. Here is a simple case: suppose that f is symmetric under some transitive group of permutations acting on the variables. Suppose also that f is odd

(i.e. $f(-x) = -f(x)$). In terms of voting methods, this means that the two candidates +1, -1 are treated equally, and that every *individual* voter is treated equally, though because we do not assume that the permutation group is 2-transitive, it is not necessarily the case that every *pair* of voters is treated equally.The entropy/influence conjecture then implies a lower bound of $I(f) \gg \log n$. An unconditional proof of this latter claim was obtained by Friedgut and Kalai using the KKL theorem mentioned earlier.For certain specific types of permutation groups, the entropy/influence conjecture predicts a precise improvement of this logarithmic lower bound. Slightly weaker versions of these improved bounds were established by Bourgain and Kalai, and
interesting to see if the techniques extend to give some weaker form of the conjecture without symmetry. Results about influence under symmetry give some hope
for similar results on influences under some sort of "spectral continuity" hypothesis. Roughly speaking, one expects assertions of the form "When the variables have geometric meaning and the Fourier coefficients respect the geometry and are "continuous", then the influence must be high."

2. **Mansour's conjecture**. Consider a Boolean function f described by a formula in conjunctive normal form of polynomial size in n. Yishay Mansour conjectured (see also this paper) that most of the Fourier coefficients are concentrated only on
polynomial number of coefficients! This conjecture is still open. Using a theorem of Hastad and Boppana one can show $I(f) \ll \log n$; thus Mansour's conjecture will follow from the entropy/influence conjecture. Mansour's motivation was to show that such Boolean functions can be statistically learned (for uniform random input) in polynomial time. This was proved by Jackson using a different method.

There are some variants and strengthenings of this conjecture:

1. **Other probability distributions**. The notions considered here and the entropy/influence-conjecture extend if we replace the unbiased Boolean variables with other product probability spaces, such as independent biased Boolean variables or Gaussian variables. It is also of interest to consider non-product spaces as well.
2. **Noise sensitivity versions**. The entropy/influence conjecture shows that the mean height of the spectrum is $\gg E(f)$. Can one guarantee a more robust conclusion, namely that the median height (or other percentiles) are also $\gg E(f)$? (This is called "noise sensitivity" but I will not explain why.) This claim is not true for the majority vote function, but some version of this may persist if one somehow rules out majority vote-like examples – for instance by considering boolean functions which are monotone, invariant under a transitive permutation group, with influence $O(n^{1/3})$ (say).
3. **Pivotals distribution versions**. If f is a boolean function and $x \in \{-1,+1\}^n$, define $\mathrm{piv}(x) := \{1 \leq k \leq n : f(\sigma_k(x)) \neq f(x)\}$. If x is selected randomly, then $\mathrm{piv}(x)$ is a random subset of $\{1,\dots,n\}$; the distribution of this random variable is called the *pivotals distribution*. Note that the expected value of this distribution is precisely I(f), so the pivotals distribution is somewhat analogous to the spectral distribution. We can ask if some version of the influence/entropy inequality applies also here. (A direct translation fails for the majority function, so we will need a weaker form.) Existence of such connection gives interesting prediction for pivotal sets for certain interesting examples.
4. **Matrix function versions**. The following related conjecture is based on discussions together with Itai Benjamini, Gadi Kozma, Elchanan Mossell, and Ofer Zeitouni.

Conjecture. There is an absolute constant $\beta > 0$ with the following property. Let $f : \mathfrak{so}(n) \to \mathbb{R}$ be an odd function on the space $\mathfrak{so}(n)$ of antisymmetric real $n \times n$ matrices, which is invariant under similarity of matrices. Then we have $I(f) \geq n^{\beta}\|f\|_2$, where the influence and L^2 norm are measured using the Gaussian probability distribution on $\mathfrak{so}(n)$. [To define the influence of non-boolean functions, one considers the average variance of these functions in each of the coordinates in turn, keeping all the other coordinates frozen.]

Kalai and Safra have a survey on issues related to these conjectures.

[*Update*, Aug 17: Maths typo corrected. – T.]

25 comments

Comments feed for this article

17 August, 2007 at 1:02 am
otto

Gil,

A small correction: At some point near the middle of your post, you start summing over S a subset of the cube, whereas it should be S a subset of {1,2, …, n}.

The same question can be asked for real-valued functions on the discrete cube, or perhaps for functions taking values in [-1,+1]. Is it essential that the range be discrete?

0 1 Rate This

Reply

17 August, 2007 at 6:19 am
John Armstrong

Hi, Dr. Kalai. What I can see looks interesting, but the vast majority of your LaTeX formulas do not parse. I'm digging through the HTML source I see on this end, but I can't see what's going wrong. Do you see them rendered properly at your end?

0 0 Rate This

Reply

17 August, 2007 at 7:22 am
Ryan O'Donnell

Otto, the range being {-1,1} is essential. Consider $f(x_1, \ldots, x_n) := (x_1 + \ldots + x_n)/n$, which has range [-1,1]. Then I(f) = 1/n, but $E(f) = (2 \log n)/n$.

Mysteriously, if you define Ent(f) to be the same as E(f) except you don't put in the "hats" (i.e., you take the entropy of f's squared values, rather than f's Fourier coefficients' squared values), then it holds that $\mathrm{Ent}(f) \leq 2I(f)$. This is the Logarithmic-Sobolev Inequality for {-1,1}^n.

Gil might also have mentioned the application of this conjecture (and related ideas) to the theory of metric space embeddings; see, e.g., Cor. 3.9 in:

Click to access nonembed-final-new.pdf

1 0 Rate This

Reply

17 August, 2007 at 7:30 am **Terence Tao**

Otto: Thanks for the typo correction! (For technical reasons, I am the one responsible for these sorts of formatting issues, though Gil is of course responsible for the content.)

John: I see the LaTeX displaying just fine; perhaps there was some temporary wordpress glitch.

0 0 Rate This

Reply

17 August, 2007 at 1:40 pm **otto**

Is there an entropic uncertainty principle saying that one of Ent(f) or Ent(\hat f) has to be large? (assuming |f| = 1)

0 0 Rate This

Reply

17 August, 2007 at 2:50 pm **Terence Tao**

Dear Otto,

There is indeed such a principle; in fact $\mathrm{Ent}(f) + \mathrm{Ent}(\hat{f})$ is minimised precisely when f is a sharp example for the discrete uncertainty principle $|\mathrm{supp}(f)||\mathrm{supp}(\hat{f})| \geq |G|$ (in this context |G| would be 2^n), or in other words when f (and hence $\hat{f}$) is a scalar multiple of a translated and modulated indicator function of a subgroup. Indeed it is not hard to see (from Jensen's inequality) that the entropy uncertainty principle implies the discrete uncertainty principle.

One particularly slick proof of the entropy uncertainty principle is to take the sharp Hausdorff-Young inequality, which with the right normalisations reads $\|f\|_{p'} \leq \|f\|_p$ for $1 \leq p \leq 2$, and differentiate (!) this inequality at $p = 2$ (noting from Plancherel's identity that we have identity at the endpoint $p = 2$).

0 0 Rate This

Reply

17 August, 2007 at 3:59 pm **Top Posts « WordPress.com**

[…] (Gil Kalai) The entropy/influence conjecture [This post is authored by Gil Kalai, who has kindly "guest blogged" this week's "open problem of the week". – […] […]

0 0 Rate This

Reply

17 August, 2007 at 5:31 pm **otto**

What about the following question (which seems morally close).

Consider $f : \{-1,1\}^n \to S^{n-1}$. Do we have

$\sum_{S \subseteq [n]} \|\hat{f}(S)\|_2^2 \log \frac{1}{\|\hat{f}(S)\|_2^2} \leq C\|f\|_{\mathrm{Lip}}$?

Even more ambitiously, if $f : \{-1,1\}^n \to R^n$, is one of
Ent(f) or Ent(\hat f) always bounded by $C\|f\|_{\mathrm{Lip}}$?
(Either concentrate of measure or concentration of spectral measure?)

0 0 Rate This

Reply

18 August, 2007 at 9:42 pm
Gil Kalai

Dear all, Greetings from Jerusalem. Thanks everybody for the interesting comments, suggestions and corrections, and thanks Terry for hosting me so helpfully.

I think that Otto's idea to find a version of the E/I-conjecture for real functions on the discrete cube is a good idea but I do not know how to do it. (As Ryan pointed out we need to modify the formulation.) Let me mention a formula by Michel Talagrand (formula (1.2) in the paper: On Russo's approximate zero-one law, Ann. of Prob. 22 (1994) 1576-1578,) which extends the KKL Theorem (in a sharp form) to arbitrary real functions on the discrete cube.

(I did not understand Otto's suggestion from remark 8 regarding functions from the discrete n-dimensional cube to S^{n-1} and R^n. Is it important that the n on both sides is the same? Also please remind me what is the LIP norm.)

The connection to metric embeddings that Ryan mentioned is indeed a very nice recent application of Fourier analysis of Boolean functions (along with a lot of other stuff). Subash Khot and Nisheeth Vishnoi made the breakthrough proving that there is a finite metric space of size n in ℓ_2^2 which requires a distortion of $(\log \log)^{1/6}$ for ℓ_1-embedding. (46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05) pp. 53-62.) Their result answered negatively a fundamental problem posed by Goemans and Linial. Along with Muli Safra we spent a lot of efforts to understand their beautiful paper. And we even have some hopes that the E/I conjecture or some relative may be relevant for possible improvements.

2 0 Rate This
Reply

19 August, 2007 at 8:30 pm
Hamed Hatami

Let me mention a problem due to Ryan O'Donnell which is related to the entropy/influence conjecture. The following statement follows immediately from the conjecture, but a valid proof that does not use the conjecture is unknown.

There exists a positive constant C such that the following holds. If $f : \{-1,1\}^n \to \{-1,1\}$, then

$\max_{S \neq \emptyset} |\widehat{f}(S)| \geq C^{-I(f)/\mathrm{Var}(f)}$.

When f is monotone, the statement follows from the KKL theorem as in that case $\widehat{f}(\{k\}) = I_k(f)$.

0 0 Rate This
Reply

14 March, 2008 at 12:09 am
Gil Kalai

I realized a typo in my presentation. In the four displaced formulas where the sums are over $S \subset \{-1,1\}^n$, it should be, of course, over $S \subset \{1,2,\ldots,n\}$.Sorry!

Here is another remark: Item 3 in the "variants" considers the distribution of the sets of pivotals, and asks if the influence is bounded below by a constant times the entropy of this pivotal distribution.

I claimed that this does not work for the balanced majority function, but it does. In this case with probability $1 - c\sqrt{n}$ we have the empty sets as the set of pivotals and otherwise the set of pivotals is a random set of size

$[n/2]$ or $[n/2]+1$. So it looks that both the influence and entropy of the pivotals distribution is $c\sqrt{n}$.

1 0 Rate This

Reply

14 May, 2008 at 7:51 am **Combinatorics, mathematics, academics, polemics, … « Combinatorics and more**

[…] and the other was about the entropy/influence conjecture. […]

0 0 Rate This

Reply

31 May, 2008 at 10:09 am **Nati's Influence « Combinatorics and more**

[…] post: The Entropy Influence Conjecture. (Terry Tao's blog) Possibly related posts: (automatically generated)Smart Mobs, Dumb […]

0 0 Rate This

Reply

31 January, 2009 at 6:36 am **Quantum boolean functions « Tobias J. Osborne's research notes**

[…] than review the theory of fourier analysis for boolean functions I'll just refer you to the post by Gil Kalai on the entropy influence conjecture which contains a very readable account. […]

0 0 Rate This

Reply

16 October, 2009 at 5:45 pm **Analysis and Probability of Boolean Functions — Gil Kalai at the UCLA-IPAM Colloquium « Not just coffee time**

[…] on Analysis and Probability of Boolean Functions. Parts of this talk have been featured in a guest blog entry over at Terence Tao's […]

0 0 Rate This

Reply

23 February, 2010 at 2:08 pm **Gil**

Mansour's conjecture for random DNF formulas was proved by Adam Klivans, Homin Lee, Andrew Wan http://eccc.hpi-web.de/report/2010/023/

1 0 Rate This

Reply

2 April, 2011 at 4:22 am **Gil Kalai**

There is a new paper on the conjecture entitled *The Fourier Entropy-Influence Conjecture for certain classes of Boolean functions*, by Ryan O'Donnell, John Wright, and Yuan Zhou, that can be found here

Click to access fei.pdf

The paper contains a proof of the conjecture for symmetric Boolean functions and in various other cases. For symmetric functions a noise sensitivity statement for the discrete directional derivatives of the function which implies the conjecture is proved.

Let me also mention that Gady Kozma asked if the conjecture remains true for functions to the unit circle in the complex plane.

0 0 Rate This

Reply

14 December, 2020 at 2:15 am **Gil Kalai**

Gideon Schechtman used analogs of the Rudin-Shapiro pairs of functions to find a counter example to the entropy/influence statement for functions to the unit complex circle. https://arxiv.org/abs/2009.12753

2 0 Rate This

Reply

---

14 June, 2011 at 12:53 pm […] theorem in terms of families of fans. (We asked about it here.)
Tentative Plans and Belated Updates II | Combinatorics and more There is a new paper on the Entropy-Influence conjecture entitled The Fourier Entropy-Influence Conjecture for certain classes of Boolean […]

0 0 Rate This

Reply

---

1 November, 2013 at 4:39 am […] post: The Entropy Influence Conjecture. (Terry Tao's […]
NatiFest is Coming | Combinatorics and more 0 0 Rate This

Reply

---

6 March, 2020 at 6:05 am […] is a 2014 post on GLL on the problem of Project 2, and here is my 2007 post on
A new PolyTCS blog! | Combinatorics and more WN on the problem of Project […]

0 0 Rate This

Reply

---

6 March, 2020 at 6:14 am

Gil Kalai

Let me mention a substantial progress and a proof of a weaker form of the conjecture in https://arxiv.org/abs/1911.10579 by Esty Kelman, Guy Kindler, Noam Lifshitz, Dor Minzer, and Muli Safra.

6 0 Rate This

Reply

---

22 March, 2020 at 2:43 am […] Let me briefly report on a remarkable new paper by Esty Kelman, Guy Kindler, Noam Lifshitz, Dor Minzer, and Muli Safra, Revisiting Bourgain-Kalai and Fourier Entropies. The paper describes substantial progress towards the Entropy-Influence conjecture, posed by Ehud Friedgut and me in 1996. (See this blog post from 2007.) […]
Kelman, Kindler, Lifshitz, Minzer, and Safra: Towards the Entropy-Influence Conjecture | Combinatorics and more

0 0 Rate This

Reply

---

29 January, 2021 at 1:03 am […] are so far three very interesting proposals there, and the first proposal is the Friedgut-Kalai Entropy/Influence conjecture. For various proposals, see also the polymath blog administered by Tim Gowers, Michael Nielsen, […]
Possible future Polymath projects (2009, 2021) | Combinatorics and more

0 0 Rate This

Reply

---

27 September, 2021 at 12:25 am […] https://www.quantamagazine.org/gil-kalais-argument-against-quantum–computers-20180207/ I think that the effort required to obtain a low enough error level for any implementation of universal quantum circuits increases exponentially with the number of qubits, and thus, quantum computers are not possible. I am pretty certain, while a little nervous to be proven wrong. Our results state that noise will corrupt the computation, and that the noisy outcomes will be very easy to simulate on a classical computer. This prediction can already be tested; you don't even need 50 qubits for that, I believe that 10 to 20 qubits will suffice. For quantum computers of the kind Google and IBM are building, when you run, as they plan to do, certain computational processes, they expect robust outcomes that are increasingly hard to simulate on a classical computer. Well, I expect very different outcomes. So I don't need to be certain, I can simply wait and see. https://terrytao.wordpress.com/2007/08/16/gil-kalai-the-entropyinfluence-conjecture/#more-41 […]
Everyday entropy: what is a map? – bonefactory

0 0 Rate This

Reply